

Apparatus for Detecting Mixed Interlaced and Progressive Original Sources in a Video Sequence

5

Background of the Invention

Field of the Invention

10 The present invention is related to the field of video signal processing and, more particularly, to methods and apparatus for identifying whether a video signal originated by means of a film source or a video camera source.

Description of the Background

15 Film and television image generation systems give rise to the appearance to the viewer of continuously moving visual images. Actually, the appearance of continuous motion results from visual and mental integration by the viewer of rapidly advancing sequences of still frame images.

20 Conventionally, in countries having a 60 hertz power grid, motion picture films are generated and are projected at one frame rate, such as 24 film frames per second, while television images are generated and displayed at another frame rate, such as 30 television frames per second.

25 Each individual frame is made up of fields. The fields are typically produced in one of two ways. The fields may be interlaced, i.e. one field being made up of even numbered lines and the other field being made up of odd numbered lines, or the fields may be progressive, i.e. one line scanned after another. For example, in the standard NTSC format used for television, the 30 frames per second are comprised of 60 interlaced fields per second, or more precisely 59.94 fields per second within the NTSC color standard signal format.

30 A video sequence may contain a mixture of interlaced video and converted progressive film material. The film material is typically sampled at 24 hertz while video is sampled at different rates: 59.94 hertz for the NTSC standard, and 50 hertz for the D1-525 and D1-625 standards. A telecine, a device for scanning and converting film into video, must maintain a proper sequence duration between the different frame rates. To do so, either a 3/2 or 2/2 pull-down technique is used to stretch the display period of film frames to the target video standard. In the 3/2 pull-down process, one film frame is used to produce three video
35 fields, and the next film frame is used to produce two fields, in a repeating 3/2 pattern. Thus,

the 3/2 pull-down is used to convert two film frames into five video fields for the D1-525 video standard. In the 2/2 pull-down system, each film frame generates two video fields, one of each type (odd/even).

When receiving video information that was originally produced on film, there is an opportunity for performing essentially error-free de-interlacing of the signal. That is because each frame of the film source is used in generating at least two video fields, representing both types (odd/even) of interlaced fields. Therefore, if a video signal can be reliably determined to have originated on film, and the video fields corresponding to a common film frame can be identified, an essentially error-free non-interlaced video frame corresponding to a single instant in time can be generated by merging the two fields. Other uses of film source identification include identification of redundant fields (which occur in 3/2 pull-down sources) to be deleted in digital transmission systems for improving channel efficiency. Also, to display video in a progressive monitor, each field needs to be interpolated vertically to the full frame size. For the film-source video, the best way to de-interlace is to merge the two fields back together to reconstruct the original progressive film frame and discard any extra repeated field.

Unfortunately, no special information is included in broadcast or other video signals to indicate which fields may have originated on film and which fields may have originated in a video camera, so the presence of film-based material must be inferred by examining differences between the luminance information of fields. That, however, can present a number of problems. For example, a strong similarity between successive video fields could indicate that they were generated from the same film frame; it could also be due to a lack of movement in the program material. Likewise, a difference between fields may indicate that the fields did not come from the same frame of information, but the difference could also be due to vertical spatial detail or transmission noise. A practical film detector must therefore distinguish between the foregoing situations.

U.S. Patent No. 5,689,301 discloses a method of film mode identification which comprises concurrently providing a first pixel from a given field and second and third vertically aligned pixels of the same horizontal position from a temporally adjacent field. The values of the pixels are compared to produce for each first pixel, a pixel difference signal having a value of zero if the value of the first pixel is intermediate the values of the second and third pixels. Otherwise, the difference signal has a value equal to the absolute value of the difference between the value of the first pixel and the value of one of the second and third pixels having a value closest to that of the first pixel. Non-zero values of the pixel difference

signals are accumulated over a predetermined portion of one field period of the video signal to provide a field difference signal. The accumulated values are then analyzed for a pattern indicative of a film source of the video signal.

Despite advances in the field of identifying original film sources, the need exists for a practical detector which can correctly detect the occurrences of film sources in video sequences, which may contain an arbitrary mix of interlaced and progressive sources.

Summary of the Present Invention

The present invention is directed to a method and apparatus of identifying the source of materials in a video sequence. A series of pseudo frames is formed, for example by interleaving, from fields in adjacent frames. A correlation value is calculated for each of the pseudo frames. The correlation value may be a sum of absolute difference (SAD) of luminance values of every neighboring scan line accumulated over the entire pseudo frame. Scene changes may be determined, for example, based on the correlation values. Frames and repeated fields are identified based on the correlation values and the scene changes. Finally, the source of each frame in the series is identified based on the identification of frames and repeated fields.

The apparatus and method of the present invention detect whether each field in a video sequence is of video (interlaced) or film (progressive) origin, and whether the field is a repeated field of the previous progressive frame. It also determines whether the field is a starting field for a new scene in the video sequence. Such information is useful for data transmission and display purposes. Those, and other advantages and benefits, will become apparent from the description of the preferred embodiment hereinbelow.

Brief Description of the Drawings

For the Present invention to be readily understood and easily practiced, the present invention will now be described, for purposes of illustration and not limitation, in conjunction with the following figures wherein:

FIG. 1 is a block diagram of an apparatus for determining the source of video sequences constructed according to the present invention;

FIG. 2 is a diagram illustrating the states of a state machine used to implement certain of the functions of the analyzer of FIG. 1;

FIG. 3 is a flow chart illustrating a method for determining the source of video sequences according to the present invention; and

FIG. 4 illustrates a system upon which the method of the present invention may be practiced.

5

Description of the Preferred Embodiments

sub a'

FIG. 1 is a block diagram of a detection apparatus 10 constructed according to the present invention for determining the source of video sequences. An external source 12 of video sequences of unknown origin provides a video sequence to the detection apparatus 10. External source 10 conforms to standard industry interfaces and provides input video sequences which may consist of an arbitrary mix of video and film source origin. The detection apparatus 10 may be operated in real time, in which case detection is done on the fly, or operated off-line.

The detection apparatus 10 is comprised of a field delay FIFO buffer 14 with a capacity of N fields. The buffer 14 typically has a minimum value for N of 4. However, if time and space are of no concern, i.e., offline non real-time systems, longer delay can be incorporated to provide more robust detection. The buffer 14 serves as a look ahead buffer for the intra-frame correlation measurements discussed next.

A circuit 16 interleaves each field with the previous field to form a pseudo frame. An intra-frame correlation (SAD) is calculated by the circuit 16 for the pseudo-frame as follows.

$$SAD = \sum_{i=0}^{Y-2} \sum_{j=0}^{X-1} |P_{ij} - P_{i+1,j}|$$

where SAD is the sum of absolute value of neighboring line differences, Y is the total number of lines in the pseudo frame, X is the total number of pixels in a line, P_{ij} is the luminance value of a pixel. The sum of the differences is a commonly used measure of correlation. Other measures can be used, including a higher power of this measure.

A circuit 18 is responsive to the intra-frame correlation SAD values. When the intra-frame correlation SAD value is higher than the previous SAD value multiplied by a pre-determined constant, a scene change is declared for the current input field. Thus:

If(SAD[1] > (SAD[2] * K))

5 Scene_change[1] = TRUE;

else

 Scene_change[1] = FALSE;

where K is a pre-determined constant. The scene change status is recorded along with the SAD value in a buffer 20.

Sub 10
~~The buffer 20 may have a capacity of N+M states, and is synchronized with the shifting of the fields within the FIFO buffer 14. The extra M states of buffering for the SAD values are required by an analyzer 22 to handle scene changes. The value of M may equal N - 1.~~

The analyzer 22 is responsive to the buffer 20. For a pseudo frame having progressive characteristics, the intra-frame SAD has a much lower value than a pseudo frame of either interlaced origin or a pseudo frame that straddles two progressive frames. That fact forms the basis for the discrimination between interlaced and progressive source videos.

The FIFO buffer 14 provides the ability to look ahead, which is necessary for the beginning of a new scene and is also used for the continuation of a scene. Accordingly, the values of SAD[i] (i from N-1 to 1) that are used by the analyzer 22 for new scenes and for continuation of scenes are P1 = SAD[N-1], P2 = SAD[N-2], P3 = SAD[N-3]. The values {P1, P2, P3} are inputs to the analyzer 22.

For the tail end of a scene, the values of SAD[i], (i from N+1 to N+M) are used. The tail-end condition is met if there is any scene change status being "TRUE" for scene_change[k], (k from 1 to N-1). In that case, P1 = SAD[N], P2 = SAD[N+1], P3 = SAD[N+2]. The values {P1, P2, P3} are inputs to the analyzer 22.

The analyzer 22 compares the values of P1, P2, and P3 as formulated above in accordance with, for example, the following equations:

If (Pi < (Pj * Kframe))

30 Is_frame = TRUE;

else

 Is_frame = FALSE;

If (Pj < (Pk * Kframe))

 Is_frame = TRUE;

else

Is_frame = FALSE;

where "Kframe" is a pre-determined constant.

5

If ((Pi, Pk "is_frame") and (Pj, Pk "is_frame"))

Is_repeated_field = TRUE;

else

Is_repeated_field = FALSE;

10

With the is_frame and is_repeated_frame determinations made, the remainder of the functions of the analyzer 22 may be implemented with, for example, a state machine implementing the following state transition table:

Current state	Condition	Next State
Interlaced	{P1, P2} "is_frame" and {P3, P2} "is_frame"	Prog_2_1
	{P1, P2, P3} "is_repeated_field"	Prog_3_1
	None of the above	Interlaced
Prog_2_1	Already pre-determined	Prog_2_2
Prog_2_2	{P1, P2} "is_frame" and {P3, P2} "is_frame"	Prog_2_1
	{P1, P2, P3} "is_repeated_field"	Prog_3_1
	None of the above	Interlaced
Prog_3_1	Already pre-determined	Prog_3_2
Prog_3_2	Already pre-determined	Prog_3_3
Prog_3_3	{P1, P2} "is_frame" and {P3, P2} "is_frame"	Prog_2_1
	{P1, P2, P3} "is_repeated_field"	Prog_3_1
	None of the above	Interlaced

15

Where:

- Interlaced means an interlaced field;
- Prog_2_1 means the first field of a progressive frame;
- Prog_2_2 means the second field of a progressive frame;
- Prog_3_1 means the first field of a repeated field progressive frame;

20

- Prog_3_2 means the second field of a repeated field progressive frame; and
- Prog_3_3 means the third field of a repeated field progressive frame.

The state transitions shown in the previous table are graphically illustrated in FIG. 2

Returning to FIG. 1, synchronized with each field 24 output from the FIFO buffer 14,

the analyzer 22 determines the output field's 24 characteristics and outputs a flag based on the next state in the state transition table indicating whether the output field 24 is a starting field of a new scene and format of the field, i.e.:

- interlaced field (interlaced), or
- The first field of a progressive frame (Prog_2_1), or
- The second field of a progressive frame (Prog_2_2), or
- The first field of a repeated field progressive frame (Prog_3_1), or
- The second field of a repeated field progressive frame (Prog_3_2), or
- The third field of a repeated field progressive frame (Prog_3_3).

The detecting apparatus 10 starts to output video fields 24, and the flags associated with each field, after the delay interposed by the FIFO buffer 14. Typical examples of video systems which can benefit from the information provided are progressive video display devices, MPEG2 video compressor, etc.

The present invention is also directed to a method of detecting mixed interlaced and progressive original sources in a video sequence. The method of the present invention is illustrated in FIG. 3. The first step of the method 26 is to buffer incoming fields. Buffering of the incoming fields is needed to select the values P1, P2 and P3 later in the process, and is also needed to enable the interleaving, at step 28, of adjacent fields to provide a pseudo frame.

After the pseudo frames have been created, through interleaving at step 28 or otherwise, an intra-frame correlation is calculated at step 30. One intra-frame correlation which may be used is based on the sum of absolute value of neighboring line differences as discussed above.

At step 32, scene changes are identified when the intra-frame correlation value is higher than a previous intra-frame correlation value multiplied by a predetermined constant. Other methods of determining when a scene changes may be used. However, because the intra-frame correlation values are available, using those values to determine scene changes is particularly advantageous.

Step 34 corresponds to the analyzer 22 of FIG. 1. In step 34, the intra-frame correlations are made available as well as the scene change information. Based on those two pieces of information, the values P1, P2 and P3 are selected and compared to one another. The basis of the comparison is a recognition that the intra-frame correlation for a pseudo frame is much lower than the intra-frame correlation of either an interlaced original or a pseudo frame that straddles two progressive frames. Based on that comparison a state machine may be used to determine the type of frame.

At step 36, frames, which are buffered at step 26, are output and synchronization with the output of information from step 34. Information may be output at step 34 in the form of flags that indicate whether the output field is a starting field of a new scene together with the format of the field.

The method of the present invention may be embodied in software and stored, for example, on a disc 38, on a computer's hard drive (not shown), or other computer readable media. The disc 38 may be used to control the operation of a computer system, such as system 40 illustrated in FIG. 4. System 40 may be comprised of a general purpose computer 42, a keyboard 44, mouse 46 and a monitor 48. Other types of input devices (scanners, microphones, etc.) and other types of output devices (speakers, printers, etc.) may be used depending upon the needs of the user. The computer 42 has a disc drive 50 for receiving the disc 38.

The present invention may also be implemented in a hardware specific implementation controlled by an application specific integrated circuit (ASIC) programmed to carry out the method as described above.

While the present invention has been described in conjunction with preferred embodiments thereof, those of ordinary skill in the art will recognize that many modifications and variations may be made. For example, other methods may be employed for creating pseudo fields. Additionally, other types of correlations may be used to generate intra-frame values for the pseudo frames. Depending upon the correlation used, the operation of the analyzer may need to be adjusted to correspond to the input information. All such modifications and variations are intended to be covered by the foregoing description and the following claims.